SEQUENCE LISTING

<110> Cannon, Paul David
      Sankuratri, Suryanarayana

<120> Human Intestinal Npt2B

<130> ROCH-001

<150> 60/119,321
<151> 1999-02-09

<160> 2

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 689
<212> PRT
<213> human

<400> 1
Met Ala Pro Trp Pro Glu Leu Gly Asp Ala Gln Pro Asn Pro Asp Lys
 1   5   10   15
Tyr Leu Glu Gly Ala Ala Gly Gln Gln Pro Thr Ala Pro Asp Lys Ser
            20   25   30
Lys Glu Thr Asn Lys Asn Asn Thr Glu Ala Pro Val Thr Lys Ile Glu
        35   40   45
Leu Leu Pro Ser Tyr Ser Thr Ala Thr Leu Ile Asp Glu Pro Thr Glu
    50   55   60
Val Asp Asp Pro Trp Asn Leu Pro Thr Leu Gln Asp Ser Gly Ile Lys
65   70   75   80
Trp Ser Glu Arg Asp Thr Lys Gly Lys Ile Leu Cys Phe Phe Gln Gly
            85   90   95
Ile Gly Arg Leu Ile Leu Leu Leu Gly Phe Leu Tyr Phe Phe Val Cys
            100   105   110
Ser Leu Asp Ile Leu Ser Ser Ala Phe Gln Leu Val Gly Gly Lys Met
        115   120   125
Ala Gly Gln Phe Phe Ser Asn Ser Ser Ile Met Ser Asn Pro Leu Leu
        130   135   140
Gly Leu Val Ile Gly Val Leu Val Thr Val Leu Val Gln Ser Ser Ser
145   150   155   160
Thr Ser Thr Ser Ile Val Val Ser Met Val Ser Ser Ser Leu Leu Thr
                165   170   175
Val Arg Ala Ala Ile Pro Ile Ile Met Gly Ala Asn Ile Gly Thr Ser
                180   185   190
Ile Thr Asn Thr Ile Val Ala Leu Met Gln Val Gly Asp Arg Ser Glu
        195   200   205
Phe Arg Arg Ala Phe Ala Gly Ala Thr Val His Asp Phe Phe Asn Trp
    210   215   220
Leu Ser Val Leu Val Leu Leu Pro Val Glu Val Ala Thr His Tyr Leu
225   230   235   240
Glu Ile Ile Thr Gln Leu Ile Val Glu Ser Phe His Phe Lys Asn Gly
                245   250   255
Glu Asp Ala Pro Asp Leu Leu Lys Val Ile Thr Lys Pro Phe Thr Lys

```
            260                 265                 270
Leu Ile Val Gln Leu Asp Lys Lys Val Ile Ser Gln Ile Ala Met Asn
    275                 280                 285
Asp Glu Lys Ala Lys Asn Lys Ser Leu Val Lys Ile Trp Cys Lys Thr
    290                 295                 300
Phe Thr Asn Lys Thr Gln Ile Asn Val Thr Val Pro Ser Thr Ala Asn
305                 310                 315                 320
Cys Thr Ser Pro Ser Leu Cys Trp Thr Asp Gly Ile Gln Asn Trp Thr
                325                 330                 335
Met Lys Asn Val Thr Tyr Lys Glu Asn Ile Ala Lys Cys Gln His Ile
                340                 345                 350
Phe Val Asn Phe His Leu Pro Asp Leu Ala Val Gly Thr Ile Leu Leu
    355                 360                 365
Ile Leu Ser Leu Leu Val Leu Cys Gly Cys Leu Ile Met Ile Val Lys
    370                 375                 380
Ile Leu Gly Ser Val Leu Lys Gly Gln Val Ala Thr Val Ile Lys Lys
385                 390                 395                 400
Thr Ile Asn Thr Asp Phe Pro Phe Pro Phe Ala Trp Leu Thr Gly Tyr
                405                 410                 415
Leu Ala Ile Leu Val Gly Ala Gly Met Thr Phe Ile Val Gln Ser Ser
                420                 425                 430
Ser Val Phe Thr Ser Ala Leu Thr Pro Leu Ile Gly Ile Gly Val Ile
                435                 440                 445
Thr Ile Glu Arg Ala Tyr Pro Leu Thr Leu Gly Ser Asn Ile Gly Thr
    450                 455                 460
Thr Thr Thr Ala Ile Leu Ala Ala Leu Ala Ser Pro Gly Asn Ala Leu
465                 470                 475                 480
Arg Ser Ser Leu Gln Ile Ala Leu Cys His Phe Phe Phe Asn Ile Ser
                485                 490                 495
Gly Ile Leu Leu Trp Tyr Pro Ile Pro Phe Thr Arg Leu Pro Ile Arg
                500                 505                 510
Met Ala Lys Gly Leu Gly Asn Ile Ser Ala Lys Tyr Arg Trp Phe Ala
                515                 520                 525
Val Phe Tyr Leu Ile Ile Phe Phe Phe Leu Ile Pro Leu Thr Val Phe
    530                 535                 540
Gly Leu Ser Leu Ala Gly Trp Arg Val Leu Val Gly Val Gly Val Pro
545                 550                 555                 560
Val Val Phe Ile Ile Ile Leu Val Leu Cys Leu Arg Leu Leu Gln Ser
                565                 570                 575
Arg Cys Pro Arg Val Leu Pro Lys Lys Leu Gln Asn Trp Asn Phe Leu
            580                 585                 590
Pro Leu Trp Met Arg Ser Leu Lys Pro Trp Asp Ala Val Val Ser Lys
    595                 600                 605
Phe Thr Gly Cys Phe Gln Met Arg Cys Cys Cys Cys Arg Val Cys
    610                 615                 620
Cys Arg Ala Cys Cys Leu Leu Cys Gly Cys Pro Lys Cys Cys Arg Cys
625                 630                 635                 640
Ser Lys Cys Cys Glu Asp Leu Glu Glu Ala Gln Glu Gly Gln Asp Val
                645                 650                 655
Pro Val Lys Ala Pro Glu Thr Phe Asp Asn Ile Thr Ile Ser Arg Glu
            660                 665                 670
Ala Gln Gly Glu Val Pro Ala Ser Asp Ser Lys Thr Glu Cys Thr Ala
    675                 680                 685
Leu
```

<210> 2

2

<400> 2

```
ctgacgtagg cccagcacct gcggaggggag cgctgaccat ggctccctgg cctgaattgg      60
gagatgccca gcccaacccc gataagtacc tcgaaggggc cgcaggtcag cagcccactg     120
cccctgataa aagcaaagag accaacaaaa ataacactga ggcacctgta accaagattg     180
aacttctgcc gtcctactcc acggctacac tgatagatga gcccactgag gtggatgacc     240
cctggaacct acccactctt caggactcgg ggatcaagtg tcagagaga gacaccaaag     300
ggaagattct ctgtttcttc caagggattg ggagattgat tttacttctc ggatttctct     360
acttttcgt gtgctccctg gatattctta gtagcgcctt ccagctggtt ggaggaaaaa     420
tggcaggaca gttcttcagc aacagctcta ttatgtccaa cccttttgttg gggctggtga     480
tcggggtgct ggtgaccgtc ttggtgcaga gctccagcac ctcaacgtcc atcgttgtca     540
gcatggtgtc ctcttcattg ctcactgttc gggctgccat ccccattatc atggggggcca     600
acattggaac gtcaatcacc aacactattg ttgcgctcat gcaggtggga gatcggagtg     660
agttcagaag agctttttgca ggagccactg tccatgactt cttcaactgg ctgtccgtgt     720
tggtgctctt gcccgtggag gtggccaccc attacctcga gatcataacc cagcttatag     780
tggagagctt ccacttcaag aatggagaag atgccccaga tcttctgaaa gtcatcacta     840
agcccttcac aaagctcatt gtccagctgg ataaaaaagt tatcagccaa attgcaatga     900
acgatgaaaa agcgaaaaac aagagtcttg tcaagatttg gtgcaaaact tttaccaaca     960
agacccagat taacgtcact gttccctcga ctgctaactg cacctcccct tccctctgtt    1020
ggacggatgg catccaaaac tggaccatga agaatgtgac ctacaaggag aacatcgcca    1080
aatgccagca tatctttgtg aatttccacc tcccggatct tgctgtgggc accatcttgc    1140
tcatactctc cctgctggtc ctctgtggtt gcctgatcat gattgtcaag atcctgggct    1200
ctgtgctcaa ggggcaggtc gccactgtca tcaagaagac catcaacact gatttcccct    1260
ttcccttgc atggttgact ggctacctgg ccatcctcgt cggggcaggc atgaccttca    1320
tcgtacagag cagctctgtg ttcacgtcgg ccttgacccc cctgattgga atcggcgtga    1380
taaccattga gagggcttat ccactcacgc tgggctccaa catcggcacc accaccaccg    1440
ccatcctggc cgccttagcc agccctggca atgcattgag gagttcactc cagatcgccc    1500
tgtgccactt tttcttcaac atctccggca tcttgctgtg gtacccgatc ccgttcactc    1560
gcctgcccat ccgcatggcc aaggggctgg gcaacatctc tgccaagtat.cgctggttcg    1620
ccgtcttcta cctgatcatc ttcttcttcc tgatcccgct gacggtgttt ggcctctcgc    1680
tggccggctg gcgggtgctg gttggtgtcg gggttcccgt cgtcttcatc atcatcctgg    1740
tactgtgcct ccgactcctg cagtctcgct gcccacgcgt cctgccgaag aaactccaga    1800
actggaactt cctgccgctg tggatgcgct cgctgaagcc ctgggatgcc gtcgtctcca    1860
agttcaccgg ctgcttccag atgcgctgct gctgctgctg ccgcgtgtgc tgccgcgcgt    1920
gctgcttgct gtgtggctgc cccaagtgct gccgctgcag caagtgctgc gaggacttgg   -1980
aggaggcgca ggaggggcag gatgtccctg tcaaggctcc tgagaccttt gataacataa    2040
ccattagcag agaggctcag ggtgaggtcc ctgcctcgga ctcaaagacc gaatgcacgg    2100
ccttgtaggg gacgcccag attgtcaggg atggggggat ggtccttgag ttttgcatgc    2160
tctcctccct cccacttctg caccctttca ccacctcgag gagatttgct ccccattagc    2220
gaatgaaatt gatgcagtcc tacctaactc gattccctttt ggcttggtgg gtaggcctgc    2280
agggcacttt tattccaacc cctggtcact cagtaatctt ttactccagg aaggcacagg    2340
atggtaccta aagagaatta gagaatgaac ctggcgggac ggatgtctaa tcctgcacct    2400
agctgggttg gtcagtagaa cctattttca gactcaaaaa ccatcttcag aaagaaaagg    2460
cccagggaag gaatgtatga gaggctctcc cagatgagga agtgtactct ctatgactat    2520
caagctcagg cctctccctt tttttaaacc aaagtctggc aaccaagagc agcagctcca    2580
tggcctcctt gccccagatc agcctgggtc aggggacata gtgtcattgt ttggaaactg    2640
cagaccacaa ggtgtgggtc tatcccactt cctagtgctc cccacattcc ccatcagggc    2700
ttcctcacgt ggacaggtgt gctagtccag gcagttcact tgcagtttcc ttgtcctcat    2760
gcttcgggga tggagccac gcctgaacta gagttcaggc tggatacatg tgctcacctg    2820
ctgctcttgt cttcctaaga gacagagagt ggggcagatg gaggagaaga aagtgaggaa    2880
tgagtagcat agcattctgc caaaagggcc ccagattctt aatttagcaa actaagaagc    2940
ccaattcaaa agcattgtgg ctaaagtcta acgctcctct cttggtcaga taacaaaagc    3000
cctccctgtt ggatcttttg aaataaaacg tgcaagttat ccaggctcgt agcctgcatg    3060
ctgccacctt gaatcccagg gagtatctgc acctggaata gctctccacc cctctctgcc    3120
```

```
tccttacttt ctgtgcaaga tgacttcctg ggttaacttc cttctttcca tccacccacc    3180
cactggaatc tctttccaaa catttttcca ttttcccaca gatgggcttt gattagctgt    3240
cctctctcca tgcctgcaaa gctccagatt tttggggaaa gctgtaccca actggactgc    3300
ccagtgaact gggatcattg agtacagtcg agcacacgtg tgtgcatggg tcaaaggggt    3360
gtgttccttc tcatcctaga tgccttctct gtgccttcca cagcctcctg cctgattaca    3420
ccactgcccc cgccccaccc tcagccatcc caattcttcc tggccagtgc gctccagcct    3480
tatctaggaa aggaggagtg ggtgtagccg tgcagcaaga ttggggcctc ccccatccca    3540
gcttctccac catcccagca agtcaggata tcagacagtc ctccctgac cctccccctt      3600
gtagatatca attcccaaac agagccaaat actctatatc tatagtcaca gccctgtaca    3660
gcattttca taagttatat agtaaatggt cttctagtgc tctcatttgg aaatgaggca      3720
ggcttcttct atgaaatgta aagaaagaaa ccactttgta tattttgtaa taccacctct    3780
gtggccatgc ctgccccgcc cactctgtat atatgtaagt taaacccggg cagggggtgt    3840
ggccgtcttt gtactctggt gattttttaga aattgaatct ttgtacttgc attgattgta    3900
taataatttt gagaccaggt ctcgctgtgt tgctcaggct ggtctcaaac tcctgagatc    3960
aagcaatccg cccacctcag cctcccaaag tgctgagatc acaggcgtga gccaccacca    4020
ggcctgattg taatttttttt ttttttttttt tttactggtt atgggaaggg agaaataaaa    4080
tcatcaaacc caaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaa                  4137
```

4

FIG. 1

```
  1   MAPWPELGDA QPNPDKYLEG AAGQQPTAPD KSKETNKNNT EAPVTKIELL

 51   PSYSTATLID EPTEVDDPWN LPTLQDSGIK WSERDTKGKI LCFFQGIGRL

101   ILLLGFLYFF VCSLDILSSA FQLVGGKMAG QFFSNSSIMS NPLLGLVIGV

151   LVTVLVQSSS TSTSIVVSMV SSSLLTVRAA IPIIMGANIG TSITNTIVAL

201   MQVGDRSEFR RAFAGATVHD FFNWLSVLVL LPVEVATHYL EIITQLIVES

251   FHFKNGEDAP DLLKVITKPF TKLIVQLDKK VISQIAMNDE KAKNKSLVKI

301   WCKTFTNKTQ INVTVPSTAN CTSPSLCWTD GIQNWTMKNV TYKENIAKCQ

351   HIFVNFHLPD LAVGTILLIL SLLVLCGCLI MIVKILGSVL KGQVATVIKK

401   TINTDFPFPF AWLTGYLAIL VGAGMTFIVQ SSSVFTSALT PLIGIGVITI

451   ERAYPLTLGS NIGTTTTAIL AALASPGNAL RSSLQIALCH FFFNISGILL

501   WYPIPFTRLP IRMAKGLGNI SAKYRWFAVF YLIIFFFLIP LTVFGLSLAG

551   WRVLVGVGVP VVFIIILVLC LRLLQSRCPR VLPKKLQNWN FLPLWMRSLK

601   PWDAVVSKFT GCFQMRCCCC CRVCCRACCL LCGCPKCCRC SKCCEDLEEA

651   QEGQDVPVKA PETFDNITIS REAQGEVPAS DSKTECTAL*  (SEQ ID NO:01)
```

BOZICEVIC, FIELD & FRANCIS LL[·]
Paula A. Borden
200 Middlefield Rd, Suite 200
Menlo Park, Ca 94025

Express Mail No. EV 036410360 US
Telephone (650) 327-3400
Fax (650) 327-3231

Inventor(s) Cannon et al.
For: "Human Intestinal Npt2B"
Docket No. ROCH-001DIV

FIG. 2i

```
   1 CTGACGTAGG CCCAGCACCT GCGGAGGGAG CGCTGACCAT GGCTCCCTGG
  51 CCTGAATTGG GAGATGCCCA GCCCAACCCC GATAAGTACC TCGAAGGGGC
 101 CGCAGGTCAG CAGCCCACTG CCCCTGATAA AAGCAAAGAG ACCAACAAAA
 151 ATAACACTGA GGCACCTGTA ACCAAGATTG AACTTCTGCC GTCCTACTCC
 201 ACGGCTACAC TGATAGATGA GCCCACTGAG GTGGATGACC CCTGGAACCT
 251 ACCCACTCTT CAGGACTCGG GGATCAAGTG GTCAGAGAGA GACACCAAAG
 301 GGAAGATTCT CTGTTTCTTC CAAGGGATTG GGAGATTGAT TTTACTTCTC
 351 GGATTTCTCT ACTTTTTCGT GTGCTCCCTG GATATTCTTA GTAGCGCCTT
 401 CCAGCTGGTT GGAGGAAAAA TGGCAGGACA GTTCTTCAGC AACAGCTCTA
 451 TTATGTCCAA CCCTTTGTTG GGGCTGGTGA TCGGGGTGCT GGTGACCGTC
 501 TTGGTGCAGA GCTCCAGCAC CTCAACGTCC ATCGTTGTCA GCATGGTGTC
 551 CTCTTCATTG CTCACTGTTC GGGCTGCCAT CCCCATTATC ATGGGGGCCA
 601 ACATTGGAAC GTCAATCACC AACACTATTG TTGCGCTCAT GCAGGTGGGA
 651 GATCGGAGTG AGTTCAGAAG AGCTTTTGCA GGAGCCACTG TCCATGACTT
 701 CTTCAACTGG CTGTCCGTGT TGGTGCTCTT GCCCGTGGAG GTGGCCACCC
 751 ATTACCTCGA GATCATAACC CAGCTTATAG TGGAGAGCTT CCACTTCAAG
 801 AATGGAGAAG ATGCCCCAGA TCTTCTGAAA GTCATCACTA AGCCCTTCAC
 851 AAAGCTCATT GTCCAGCTGG ATAAAAAAGT TATCAGCCAA ATTGCAATGA
 901 ACGATGAAAA AGCGAAAAAC AAGAGTCTTG TCAAGATTTG GTGCAAAACT
 951 TTTACCAACA AGACCCAGAT TAACGTCACT GTTCCCTCGA CTGCTAACTG
1001 CACCTCCCCT TCCCTCTGTT GGACGGATGG CATCCAAAAC TGGACCATGA
1051 AGAATGTGAC CTACAAGGAG AACATCGCCA AATGCCAGCA TATCTTTGTG
1101 AATTTCCACC TCCCGGATCT TGCTGTGGGC ACCATCTTGC TCATACTCTC
1151 CCTGCTGGTC CTCTGTGGTT GCCTGATCAT GATTGTCAAG ATCCTGGGCT
1201 CTGTGCTCAA GGGGCAGGTC GCCACTGTCA TCAAGAAGAC CATCAACACT
1251 GATTTCCCCT TTCCCTTTGC ATGGTTGACT GGCTACCTGG CCATCCTCGT
1301 CGGGGCAGGC ATGACCTTCA TCGTACAGAG CAGCTCTGTG TTCACGTCGG
1351 CCTTGACCCC CCTGATTGGA ATCGGCGTGA TAACCATTGA GAGGGCTTAT
1401 CCACTCACGC TGGGCTCCAA CATCGGCACC ACCACCACCG CCATCCTGGC
1451 CGCCTTAGCC AGCCCTGGCA ATGCATTGAG GAGTTCACTC CAGATCGCCC
1501 TGTGCCACTT TTTCTTCAAC ATCTCCGGCA TCTTGCTGTG GTACCCGATC
1551 CCGTTCACTC GCCTGCCCAT CCGCATGGCC AAGGGGCTGG GCAACATCTC
1601 TGCCAAGTAT CGCTGGTTCG CCGTCTTCTA CCTGATCATC TTCTTCTTCC
1651 TGATCCCGCT GACGGTGTTT GGCCTCTCGC TGGCCGGCTG GCGGGTGCTG
1701 GTTGGTGTCG GGGTTCCCGT CGTCTTCATC ATCATCCTGG TACTGTGCCT
1751 CCGACTCCTG CAGTCTCGCT GCCCACGCGT CCTGCCGAAG AAACTCCAGA
1801 ACTGGAACTT CCTGCCGCTG TGGATGCGCT CGCTGAAGCC CTGGGATGCC
1851 GTCGTCTCCA AGTTCACCGG CTGCTTCCAG ATGCGCTGCT GCTGCTGCTG
1901 CCGCGTGTGC TGCCGCGCGT GCTGCTTGCT GTGTGGCTGC CCCAAGTGCT
1951 GCCGCTGCAG CAAGTGCTGC GAGGACTTGG AGGAGGCGCA GGAGGGGCAG
2001 GATGTCCCTG TCAAGGCTCC TGAGACCTTT GATAACATAA CCATTAGCAG
2051 AGAGGCTCAG GGTGAGGTCC CTGCCTCGGA CTCAAAGACC GAATGCACGG
2101 CCTTGTAGGG GACGCCCCAG ATTGTCAGGG ATGGGGGGAT GGTCCTTGAG
2151 TTTTGCATGC TCTCCTCCCT CCCACTTCTG CACCCTTTCA CCACCTCGAG
2201 GAGATTTGCT CCCCATTAGC GAATGAAATT GATGCAGTCC TACCTAACTC
2251 GATTCCCTTT GGCTTGGTGG GTAGGCCTGC AGGGCACTTT TATTCCAACC
2301 CCTGGTCACT CAGTAATCTT TTACTCCAGG AAGGCACAGG ATGGTACCTA
2351 AAGAGAATTA GAGAATGAAC CTGGCGGGAC GGATGTCTAA TCCTGCACCT
2401 AGCTGGGTTG GTCAGTAGAA CCTATTTTCA GACTCAAAAA CCATCTTCAG
2451 AAAGAAAAGG CCCAGGGAAG GAATGTATGA GAGGCTCTCC CAGATGAGGA
2501 AGTGTACTCT CTATGACTAT CAAGCTCAGG CCTCTCCCTT TTTTTAAACC
2551 AAAGTCTGGC AACCAAGAGC AGCAGCTCCA TGGCCTCCTT GCCCCAGATC
2601 AGCCTGGGTC AGGGGACATA GTGTCATTGT TTGGAAACTG CAGACCACAA
```

FIG. 2ii

```
2651 GGTGTGGGTC TATCCCACTT CCTAGTGCTC CCCACATTCC CCATCAGGGC
2701 TTCCTCACGT GGACAGGTGT GCTAGTCCAG GCAGTTCACT TGCAGTTTCC
2751 TTGTCCTCAT GCTTCGGGGA TGGGAGCCAC GCCTGAACTA GAGTTCAGGC
2801 TGGATACATG TGCTCACCTG CTGCTCTTGT CTTCCTAAGA GACAGAGAGT
2851 GGGGCAGATG GAGGAGAAGA AAGTGAGGAA TGAGTAGCAT AGCATTCTGC
2901 CAAAAGGGCC CCAGATTCTT AATTTAGCAA ACTAAGAAGC CCAATTCAAA
2951 AGCATTGTGG CTAAAGTCTA ACGCTCCTCT CTTGGTCAGA TAACAAAAGC
3001 CCTCCCTGTT GGATCTTTTG AAATAAAACG TGCAAGTTAT CCAGGCTCGT
3051 AGCCTGCATG CTGCCACCTT GAATCCCAGG GAGTATCTGC ACCTGGAATA
3101 GCTCTCCACC CCTCTCTGCC TCCTTACTTT CTGTGCAAGA TGACTTCCTG
3151 GGTTAACTTC CTTCTTTCCA TCCACCCACC CACTGGAATC TCTTTCCAAA
3201 CATTTTTCCA TTTTCCCACA GATGGGCTTT GATTAGCTGT CCTCTCTCCA
3251 TGCCTGCAAA GCTCCAGATT TTTGGGGAAA GCTGTACCCA ACTGGACTGC
3301 CCAGTGAACT GGGATCATTG AGTACAGTCG AGCACACGTG TGTGCATGGG
3351 TCAAAGGGGT GTGTTCCTTC TCATCCTAGA TGCCTTCTCT GTGCCTTCCA
3401 CAGCCTCCTG CCTGATTACA CCACTGCCCC CGCCCCACCC TCAGCCATCC
3451 CAATTCTTCC TGGCCAGTGC GCTCCAGCCT TATCTAGGAA AGGAGGAGTG
3501 GGTGTAGCCG TGCAGCAAGA TTGGGGCCTC CCCCATCCCA GCTTCTCCAC
3551 CATCCCAGCA AGTCAGGATA TCAGACAGTC CTCCCCTGAC CCTCCCCCTT
3601 GTAGATATCA ATTCCCAAAC AGAGCCAAAT ACTCTATATC TATAGTCACA
3651 GCCCTGTACA GCATTTTTCA TAAGTTATAT AGTAAATGGT CTTCTAGTGC
3701 TCTCATTTGG AAATGAGGCA GGCTTCTTCT ATGAAATGTA AAGAAAGAAA
3751 CCACTTTGTA TATTTTGTAA TACCACCTCT GTGGCCATGC CTGCCCCGCC
3801 CACTCTGTAT ATATGTAAGT TAAACCCGGG CAGGGGCTGT GGCCGTCTTT
3851 GTACTCTGGT GATTTTTAGA AATTGAATCT TTGTACTTGC ATTGATTGTA
3901 TAATAATTTT GAGACCAGGT CTCGCTGTGT TGCTCAGGCT GGTCTCAAAC
3951 TCCTGAGATC AAGCAATCCG CCCACCTCAG CCTCCCAAAG TGCTGAGATC
4001 ACAGGCGTGA GCCACCACCA GGCCTGATTG TAATTTTTTT TTTTTTTTTT
4051 TTTACTGGTT ATGGGAAGGG AGAAATAAAA TCATCAAACC CAAAAAAAAA
4101 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAA    (SEQ ID NO:02)
```